Toward an Analytical Framework for the Study of Distributed Problem-Solving Networks

A Perspective on the OII-MTI Project

By

Paul David

27 August 2007

The *DPSN* project is about "network enabled" distributed problem-solving organizations. In this context to be network enabled means that the emergence of the organization and its sustained functioning depends upon the use to computer-mediated telecommunication networks in recruiting, mobilizing and coordinating spatially distributed knowledge resources.

An *organization* is a multi-component entity within which there are defined structural relationships and distinct roles among the constituent components or participating "actors", and means of coordinating their actions and interactions so as to produce coherent outcomes that link the entity to a supporting environment.

The foregoing generic and necessarily abstract conceptualization subsumes both *social organizations*, and machine organizations, and mixtures thereof. Social organizations are ensembles of volitional agents, and may be the result of deliberate and goal-oriented design; or may be an emergent product of a self-organizing system of interactions among its constituent actors. Human social organizations are of both kinds, and mixed human-machine organizations necessarily contain elements of deliberate design. As a rule, goal-oriented organizations that have been purposefully constructed will have some manifest functions, but may also possess "latent functionalities" that are important in sustaining both the participation of the volitional actors and stabilizing the structure of relationships among them—thereby permitting the organizational form, or system, to reproduce itself through time.

Knowledge resources here refers to human capabilities, and is to be distinguished from information, the latter being "structured data." Knowledge is a capacity for structuring data, encoding it as information for storage, retrieval, and processing – including the "creative" generation of novel information, and its transmission to other agents that have the capacity to read and interpret its code and extract from it signals upon which actions (including maintenance of the status quo) are based. The class of organizations which our project is particularly concerned are "information-intensive", in the sense that make extensive use of information (along with inputs of tangible resources and services) to produce information.

The project descriptor and its acronym (DPSN) can therefore be read as conflating "organization" in the foregoing sense with "network," by employing the latter term both literally and metaphorically. Technical communication networks are enabling information tools for social organizations that are not spatially co-located; but decentralized systems of peer-to-peer interaction are often referred to

(metaphorically) as "network organizations" and contrasted with centrally managed and hierarchically structured organizational forms that are more characteristic of business corporations and public administrative bureaucracies.

"Network organization" is thus a term that can apply both to social organizations in which the relationships among the constituent actors are structured in some part by formal contracts, and to "communities" in which peer-to-peer interactions are not contractually based. In the absence of complete (fully specified) contracts, or, indeed, any contracts, social interactions within organizations will be governed by informal but mutually acknowledged norms of behavior with which compliance is mutually expected. Many of the more familiar large, goal-directed human social organizations coordinate their internal workings by using mixtures of (incomplete) contracts and informal social norms.

Part I

A Simple and Generic Conceptual Framework for DSPN Studies

The DPSN project is studying a group of "cases" that have been selected as examples of spatially distributed and information-intensive organizations that have emerged with the evolution of the Internet. The accompanying "tableau" of projected case studies (Table 1/v4/27.08.07) presents them in two groupings. Group I contains a variety of distributed "communities" that have established identities associated with the creation and production of particular technical and cultural artifacts or "products". These offer an opportunity to examine problem-solving procedures and routines that the community has utilizes to address (and solve) "problems" whose resolution is perceived to be a critical task for the achievement of its functional purpose(s). The cases in Group II are focused on Internet enable "platform mechanisms" that are deployed by information-intermediaries: they gather data and information from distributed individual actors and provide a means of processing ("mining") it to extract "signals" that offer solutions to "problems" that are of interest to third-party organizations.

Within each of those two groupings may be found a number of case studies concerned with organizations and mechanisms that have emerged quite recently into visibility on the Internet, and others that have evolved with repeated use over considerably longer time periods. In a broad sense, interest in the results of examining of the former, derives from questions about the viability and potentialities for long-term survival of these novel organizational forms and practices; whereas, for the case studies of more established organizations, there is an interest in the extent to which the contextual requirements for the particular solution-routines that they have been using, as well as their performance attributes, are such that they might be successfully adapted to other quite different domains of application. In this regard, the design of the project affords an opportunity to examine two aspects of the performance of contemporary DPSNs and may carry important implications for the future organization of information-intensive production activities on the Internet.

All the cases in this project are concerned with a problem-solving "organization" or system, whether or not the studies that can be carried out within

this exploratory phase are able to examine all the parts of the system in equal detail. It is necessary at a minimum to relate the focus of all the case studies to the respective larger "system" of which it is a part (if not the whole), simply for the purposes of understanding the performance criteria and requirements that have a bearing on the contemporary and future domains of applicability.

To speak of such a system or organization as logically complete, one therefore must be able to identify its four essential aspects or dimensions: the "problem-holder(s)", the "problem-solvers", the "communication structure" for information flows among the problem-holder(s) and problem-solvers, and the nature of "the problem(s)" that are being address and solved (at least in some degree). To describe each case in reference to those four dimensions consequently provides a basic foundation for recognizing aspects of resemblance and differentiation among them.

Further, by articulating those four dimensions, and examining each of the cases in reference to them, it is possible to assemble the empirical ingredients for analyses that will shed light on why these particular DPSN arrangements have emerged and where and how they may be replicated. There is of course more than one theoretical perspective on organizational behavior that could be illuminating for this purpose, and the following brief remarks on the view from economics is not meant to preclude other approaches that might be explored.

To keep the analysis at its simplest level, one can start from the observation that *the demand side* of the story for a particular mode of *distributed* problem-solving is likely to reside in the characteristics of problem-holders and the attractiveness for them of seeking to mobilize external distributed knowledge resources, rather than assign the problem-solution task to an internal co-located team, or "out-source" it to another organization. "Attractiveness" from the perspective of the problem-holder, is a matter of comparative efficacy (including speed), cost, reliability, and control over information about the nature of the organization's problem and whatever solution may be obtained.

The nature of the problems that problem-holders select for solutions that are provided by external (distributed) actors, in general, will not be a representative sample of all those that arise in the course of their activities. Rather, the selectivity bias will reflect the problem-holder's capabilities with regard to the kinds of problems that can readily be formulated in ways that allow external agents (without extensive knowledge of the problem-holder's purposes, plans and internal organizational practices) to work effectively upon feasible solutions. It will, in addition, reflect the ability of the problem-holder to screen and evaluate the comparative merits ("worth") of the solutions that would be presented by external "solvers," and thus to select the best from among those received.

Turning to consider the *supply side* of the story, one has to recognize at least two possible arrangements: the problem holder can deal directly with actors that are the distributed sources of knowledge capabilities that can be brought to work on solving the whole or parts of "the problem"; or there may be an intermediating information agency, such as a web-based platform provider, that performs the function of mobilizing the sources of information from which proposed solutions can be offered to a problem-holder. In the former situation, the problem-holding entity

has to consider the costs of identifying, contacting and transacting with a selected population of problem-solvers – identifying the size and the array of capabilities among them that are likely to be required to generate an acceptable solution to the chosen problem in a timely fashion. The process they use, moreover, has to be one whose nature and implementation is consistent with the continued participation of a sufficiently large and competent problem-solving community; otherwise the mechanism will not be sustainable. One way of putting this is to say that it is necessary, in order for the system to be in equilibrium (and therefore reproduce itself through time), that the expectations of the suppliers of the solutions and those of the problem-holder(s) be consistent and satisfied -- directly or indirectly -- by the distribution of outcomes. We can in this way speak about the matching of the "supply" and "demand" for problem solving without thereby suggesting that this is achieved through market institutions that commodify "solutions" and set prices so as to "clear the market."

In situations where there is an intermediating agency involved, it is possible for that entity to be specialized in collecting and mining information of more or less pre-specified kinds from populations of distributed sources that are "closed" or "open", and to offer such "products" as potential solutions to problem-holders. Platform operating intermediaries will perform more efficiently when they are able to develop relationships with "clients" that enable them to anticipate the array of "problems" for which solutions are in strong demand. At the same time, they will need to provide incentives or some forms of satisfaction that are sufficient to maintain the participation of the sources from which they gather data and information for analysis. For problem holders, the attractiveness of transacting for information with intermediaries will be influenced by the same array of considerations that would apply were they to directly manage their interactions with the distributed information sources, with a further complication that is called a "common agency problem": the principal's potential "costs" may also include those assignable to the expected risks entailed in sharing information about the nature of their problems with an agency that also may use that knowledge to serve the interests of its own (other) clients.

Whatever arrangements exist for bringing the needs of problem-holders together with the capabilities of distributed problem-solvers, any specific problem-solving system or "organization" will be characterized a particular *structure of communications.* This structure will determine – by technical or other means -- when, and in what regards the problem-solvers are permitted to exchange information with one another about the problems on which they are working. Similarly, it will specify the way that information flows pass between problem-solvers and problem holders. The controls placed on flows of information within the system – whether by technological constraints or by contractual or other regulations -- are likely to have a critical bearing upon the efficiency of the problem-solution process, its robustness to exogenous disruptions, and the distribution among the participating actors of the social benefits that result from the process. Being able to fully describe the permitted patterns of information flows within the organization as a whole, therefore, is likely to be an important step toward understanding such a system's potential functional performance, and the ways in which its realized performance will be evaluated from the different private vantage points of the parties that are involved.

28 August 2007

Table 1. Evolving "Tableau" of Prospective Case Studies and Teams

Type of Case	Description of problem-holder's activity	Economic sector or industry	Problem Solution-Products	Case study	Draft by OII Researcher	Collaborating researchers
Complex and Novel Problems –Addressed by organisation of "problem-holding problem-solvers"	Fundamental science: multi- team research coordination	Public sector science	Design of large-scale HEP detector facility	Atlas-CERN	**Paul David**	Philipp Tuertscher; Felix Reed-Tsochas and others?
	Open Source Software:reliability maintenance	IT Software	Web browser	Mozilla Bug-patching	Matthijs den Besten	J-M Dalle; **Paul David**
	Peer-production of free instructional material	Publishing	Text Books	FHSST—High school science, math texts (.Cape Town, S.A.)	Max Loubser	Matthijs den Besten; **Paul David**
	Peer-Production of free compendia content : governance	Encyclopedia Publishing	Process to manage conflicts over open content	Resolving problem of a contentious entry in Wikipedia	Max Loubser	Matthijs den Besten; **Paul David**
	Peer-Production of free compendia content: quality assurance	Encyclopedia Publishing	Readability standardisation of entries in encyclopedia	"Simple Wikipedia"	Matthijs den Besten	J-M Dalle; Max Loubser; **Paul David**
	? Open Content Film Production projects in Europe	Entertainment	?Community-based financing and creation of new films using open content	? "Swarm of Angels" (in Brighton, UK); "Cinema 2.0" (in Turin, Italy)	?Aldo Geuna	? Irene Cassarino; ? Wolf Richter; **? Paul David**;

Type	Description of problem-solving mechanism	Economic sector or industry	Problem Solution-Products	Case study	Draft by OII Researcher	Collaborating researchers
Platforms for generation and miming of data from distributed agents' social interactions	Information exchange on closed platform	Clinical healthcare, Pharmacology	"Threads": e.g., drug side-effects, off-label uses, epidemics	Sermo	David Bray	Karen Croxson; **Bill Dutton**
	Multi-player Game Environments	Enterprise solutions	Prioritised messages, rationalised intra-organisational communications	Seriosity	David Bray	Karen Croxson; **Bill Dutton**
	News aggregation platforms	Publishing; public opinion research	News production, filtering/attention focus	Digg News / Reddit?	Wolf Richter	David Bray; **Bill Dutton**
	Prediction Markets: for public events	Public opinion research	Tracking evolution of public opinion	Iowa election markets; Betfair(sports)	Karen Croxson	**Bill Dutton**
	Prediction Markets:internal corporate events	Enterprise solutions	Aggregation of intra-company information on business process status	Google &/or HP (?) product-launch betting	Karen Croxson	**Bill Dutton**
Special Cross-cutting Topics	Webmetric analysis approaches				Rob Ackland	**Bill Dutton**
	Survey Approaches				**Bill Dutton**	David Bray
	Intellectual Property				Wolf Richter	**Paul David**
	?Motivations in Design and Practice				?Karen Croxson	**?Bill Dutton**

Part II

Elaborating an Analytical Taxonomy for Case Studies of DPSN*s*

Four basic structural of DPSN organizations that have been identified in the discussion of Part I:

    I. ***The Problem(s****)* for which answers are provided -- their nature and formulation
    II. ***The Problem-holder(s***)-- their organizational characteristics and capabilities
    III. ***The Problem-solvers*** -- their capabilities, motivations and situations
    IV. ***The Communication Structure*** -- among and between problem-solvers and – holder(s)

Taken together these form the main pillars upon which descriptions and analyses of particular cases can be constructed, in terms that permit a comparisons among them within a unifying framework that can be expanded to absorb findings from the study of other examples of the general phenomenon.

Many aspects of structural similarity and difference among specific "cases" can be identified by a taxonomic unfolding, or "unpacking" of each of the foregoing major dimensions. That can be done by drawing a number of distinctions in regard to each "dimension," selecting alternatives among those that are likely to have some important causal relationship to the actors' behaviors-- whether as the endogenous resultants of decisions, or as exogenous structural features that are among the factors determining such decisions.

In principle, such an "unfolding" exercise could yield a very large number of possible "types" of DPSNs, each occupying a distinct position in a multi-dimensional space. But, while that might be of theoretical interest, it is neither necessary nor practical for our work to undertake and exhaustive mapping of that entire space *ex ante*. Some of its regions might well be empty of any real world examples, and much of it will not apply to those cases that have been selected in this exploratory project. Indeed, in our exploratory project we are only preparing a few of the conceivable number of distinctive arrangements that might be available for study. So, rather than go in that direction, we shall proceed from the outset to consider only a limited number of categorical distinctions that might turn out to be pertinent in regard to each of the four basic dimensions, and allow for further, more refined distinctions to be made if and where it appears important in the context of one or more specific case studies.

A straightforward way to start this "unfolding" operation is to work through a logical succession of questions under each of the four main (structural) headings, beginning with gross structural distinctions that can be applied across all the cases. The particular answer to each of these questions may be of greater importance in some cases than in others, and where it does seem to be potentially crucial, further differentiations and their implications can be, and should be pursued in the presentation of the case details.

The following set of questions has been worked out in discussions with Wolf Richter, who undertook to test the applicability of a basic set of questions to one of the cases in Group II of Figure 1's tableau.

I. The Nature of the Problem

1. What is the nature of "the problems" that the problem-holder(s) need to have solved, or that the problem-solving agents are solved for them?
- Are they related to innovative projects per se, or do they concern the performance of routine organizational operations?
- Are they about something that is objective in nature, or subjective: i.e., does it require technical and/or scientific expertise relating to product designs and methods of production, or does it involve market research, or public opinion research about the current state of opinions and preferences?
- Are the multi-part, requiring solutions to many sub-problems or do they involve an integral question?
- Are they *complex*, in the sense of having not only multiple parts or tasks to be solved, but a having tasks that are *not semi-decomposable* – so that the solution of one task will affect the solutions to a number of other tasks that cannot be solved in parallel or grouped into in a sequential solution order?

II. The Problem holders

2. Who is (are) the problem-holder(s)?
- A centrally managed entity? A consortium of organizations? A collectivity of individuals?
- Is the problem-holding entity distinct from the problem-solvers? {see Q. 5ff below}
- If the problem-holder is an a commercial enterprise, its it's line of business marketing goods and services over which it has some "market power", or it a price taker in a competitive market – so that it is indifferent whether competitors do or don't know about its production methods?
- If a business firm, does its market setting and business model require that it have exclusive intellectual property rights over its product design or production methods?

3. What are the problem holder's own internal organizational capabilities for structuring and addressing questions as problems?
- Can problem holders pre-structure their problem for external expert solution or do they need to see options which they can recognize as solving some problem(s) they had not focused upon?
- Do they have one or more competent internal problem-solving organizations, so that turning to external problem solvers could be viewed as a "buy" vs. make" decision?
- Are they able to structure the problem(s) so that they can organize a distributed external approach to the solution(s) by distributed knowledge resource providers?   {See answer to Q.1, bullets 3, 4}
- What are the relevant decision criteria "make or buy" decisions in the specific situation of problem-holders?

4. Has the problem-holder(s) organized the solution process themselves and provided explicit incentives for problem-solvers, or does the problem-holder transact indirectly with the latter, through the services it obtains from an external intermediary?

- If there is an external intermediary, is this agency an enterprise providing a platform and gathering solution options from many distributed sources?
- Alternatively, is the external agent a closed and hierarchically managed research enterprise with which problem solvers may contract?
- Or, are the external agents members of private professional or personal networks that can be accessed directly and informally for reciprocated exchanges of expert knowledge by a problem-holder (or individual employees)?
- Are "rewards" of any kind announced publicly – whether by the problem-holder directly, or indirectly via an intermediary agent -- for the solutions of pre-specified problems? If so, do these take a direct or an indirect pecuniary form, and do they stipulate precise requirements for a solution that will qualify for the reward? If not, is a non-pecuniary reward—such as public recognition—explicitly promised for successful solutions?
- Is the manner in which successful solutions (those qualifying for explicit rewards) of will be selected announced publicly to prospective problem-solvers?

III. Problem Solvers

5. Who are the problem-solvers, and what are the motivating incentives and expectations of (and uncertainties of the surrounding distribution of) "reward" that the solution process offers for those who are contributing to it as problem-solvers, or as intermediating (broker) agents?

- Is membership in the set of problem-solvers open, or closed; and if closed, on what principle and by what means?
- Are the problem-solvers independent experts, i.e. "soloists"?
- If they are expert, are they an ensemble of players, i.e., a "community" that has formed and continues to perform without a central conductor? or are they members of an "orchestra" or "big band", i.e., contractually related to a corporate entity and coordinated by a "conductor" (or "band-leader") and "first violinist" (or" lead instrumentalist") the counter parts for the less musically inclined being company directors and managers
- What is the form of the dominant effective reward set for the problem solvers, and who sets it if it is pecuniary, the problem holder, an intermediary agent, or the collectivity of problem solvers?

IV. Structure of Communications

6. What is the permitted/intended structure of communications between problem solvers and problem holders under the specific arrangements in the case?

- Do problem holders and problem solvers interact directly or facilitated by an intermediary?

- Who can contact who at what stages, with what kinds of messages?
- What is the actual pattern of communications among problem-solvers, and between them and the problem holder(s) -- subject to the constraints imposed by the solution mechanism that is being used?
- Are the contents of information flows archived for subsequent retrieval and analysis, and if so by whom?

7. What is the permitted structure of communication among problem holders and problem solvers?
- Is all communication facilitated by an intermediary or is there direct interaction among individual problem holders and problem solvers?
- Are problem holders / solvers communicating regularly outside the scope of the particular problem to solve, i.e. because they belong to a larger community or has a communication structure been established *ad hoc* for the purpose of solving the problem?

A Cross-cutting question on Intellectual Property Rights arrangements:

8. **When solution options are obtained, who holds the legal rights to exploit them (as trade secrets or IP), and who holds the rights when particular solution options are selected by the problem holder(s)?**
   - How do these arrangements affect the flow of communications in the solution process?
   - How do these arrangements affect the payoffs distribution of the problem-solvers and the problem holders?

Part III

Brief Notes on Implementation of the Proposed Framework

The main purpose of the foregoing questions is to provide some common conceptual tools and terminology that (1) facilitate comparative discussion of the selected cases, (2) highlight structural features relevant to the specification of performance metrics, (3) relate these to other studies of distributed problem-solving.

A further purpose is to suggest the array of questions that researchers responsible for preparing drafts describing the "case" should (eventually) be able to answer. But it is not proposed that the descriptive material should be structured as responses to these questions, or that it would address them in the particular order in which they have been set out in the preceding pages.

Quite the contrary, it is deliberately left for those who have familiarized themselves with the details of the case, including a knowledge of the circumstances of the emergence and evolution of the problem-solving organization, to find and point out significant connections among the issues identified by particular groups of questions that may make the "logic" of its features and performance properties more intelligible.